



並列コンピュータシステムの更新について

著者	小林 広明, 岡部 公起
雑誌名	SENAC : 東北大学大型計算機センター広報
巻	39
号	1
ページ	1-5
発行年	2006-01
URL	http://hdl.handle.net/10097/00124375

並列コンピュータシステムの更新について

情報シナジーセンタースーパーコンピューティング研究部

小林広明、岡部公起

1. はじめに

東北大学情報シナジーセンター(以下、センター)では、大規模科学計算システムのうち、並列コンピュータシステム TX7/AzusA を、スカラ並列型とベクトル並列型のハイブリッド構成の並列コンピュータシステムに更新し、来年 3 月から新システムにて運用を開始します。本稿では、新しい並列コンピュータシステムの紹介、および移行日程の概略について述べます。

2. システム構成

センターでは、現在の大規模科学計算システムの利用状況から、新並列コンピュータシステムに求められる性能・機能を検討し、仕様策定を行ってきました。現在、本センターは、スーパーコンピュータとしてベクトル型スーパーコンピュータ SX-7(日本電気(株)製、232CPU、1792GB、2048GFLOPS)、並列コンピュータとして TX7/AzusA(日本電気(株)製、112CPU、176GB、358GFLOPS)を運用しています。平成 15 年 1 月に導入したスーパーコンピュータ上では高ベクトル化率のプログラムが常時実行されており、演算性能と主記憶容量の格段の強化を行ったにもかかわらず、計算需要はすでに処理能力の限界を超えており、増強の必要に迫られています。また、平成 14 年 1 月に導入した並列コンピュータについても、運用開始からすでに 4 年近くになり、演算性能と主記憶容量の不足は深刻であり、計算モデルの高精度化・大規模化を推進するためには、演算性能と主記憶容量を大幅に強化する必要があります。このことから、新並列コンピュータシステムは、スカラ並列処理性能に優れた演算サーバとベクトル並列処理性能に優れた演算サーバから構成されるハイブリッド型並列コンピュータシステムを導入することにしました(図 1 参照)。

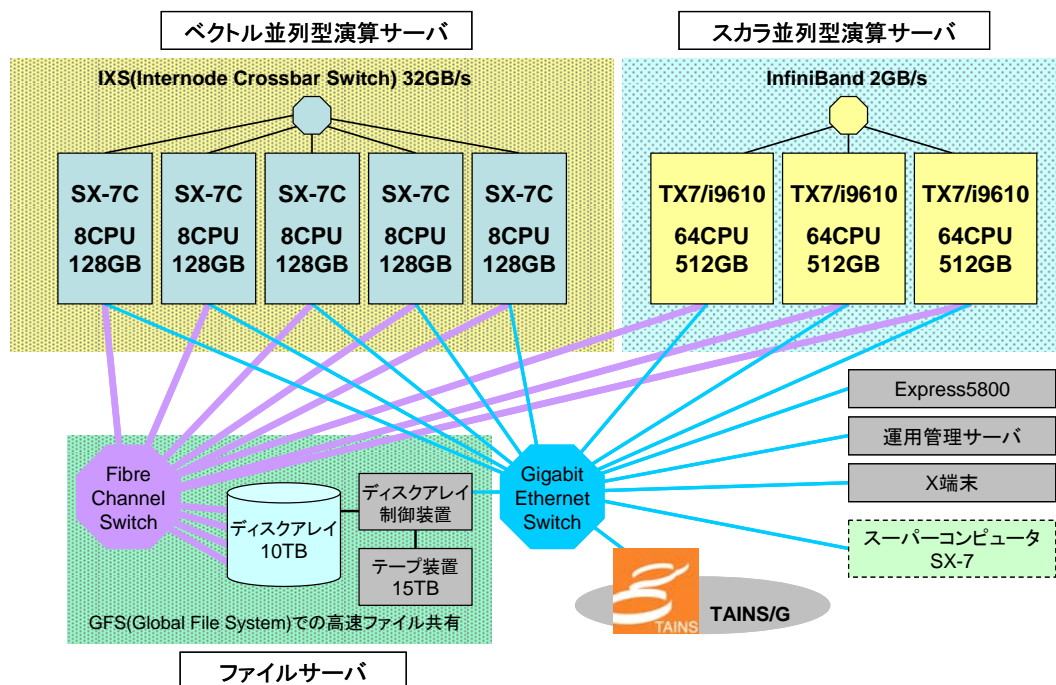


図 1 新並列コンピュータシステム

(1) スカラ並列型演算サーバ(TX7/i9610、日本電気株製)

スカラ並列型演算サーバは 3 つのノードから構成され、それぞれのノードは、Intel の最新の Itanium2 プロセッサ 64CPU と 512GB の主記憶容量を持ち、Linux のもとで動作します。Itanium2 プロセッサは 1.6GHz で動作し、ノードあたりの最大演算性能は 409.6GFLOPS となります。したがって、スカラ並列型演算サーバの CPU あたりの性能は、現在運用中の TX7/AzusA の CPU の性能の 2 倍の演算性能を持つことになり、表 1 に示す本サーバで実行される Gaussian などアプリケーションプログラムの高速化に加えて、ベクトル演算に不向きなプログラムや並列処理の不可能なプログラムの高速実行も期待できます。さらに、ノード間は 2GB/s の双方向通信性能を有する InfiniBand で接続されており、MPI などを用いた複数のノードを利用した並列処理においても、スケラブルな性能を発揮することができます。

表 1 スカラ並列型演算サーバで利用可能なアプリケーション

分子軌道計算ソフトウェア	Gaussian03
統合型数値計算ソフトウェア	MATHEMATICA
汎用構造解析プログラム	MSC.MARC MSC.MARC Mentat

またスカラ並列型演算サーバは、並列プログラムの開発、および既存のベクトル並列型スーパーコンピュータ SX と次に述べるベクトル並列型演算サーバのフロントエンドとしての役割も担っており、表 2 に示すコンパイラおよびライブラリ等が整備されています。

表 2 スカラ並列型演算サーバで利用可能なコンパイラおよびライブラリ

FORTRAN95	ISO/IEC 1539-1:1997,自動並列化, OpenMP
C++	ISO/IEC 14882:1998,自動並列化, OpenMP
MPI/EX	TX7 用メッセージパッシングライブラリ
FORTRAN90/SX	ISO/IEC 1539-1:1997,SX-7/SX-7C 用クロスコンパイラ
C++/SX	ISO/IEC 14882:1998,SX-7/SX-7C 用クロスコンパイラ
MPI/SX	SX-7/SX-7C 用メッセージパッシングクロスライブラリ
ASL	数値計算ライブラリ

なお、Itanium2 に対応していないアプリケーション MATLAB(対話型解析ソフトウェア)と SAS(統計解析システム)については、Express5800(Intel Xeon 4 プロセッサ、3.66GHz)を用いてサービスを提供します。

(2) ベクトル並列型演算サーバ(SX-7C、日本電気㈱製)

ベクトル並列型演算サーバは、現在運用中のベクトル並列型スーパーコンピュータ SX-7 の過負荷状態を緩和するために導入するもので、演算周波数 2GHz で動作し、CPU あたり 16GFLOPS(最大ベクトル性能)、ノードあたり 128GFLOPS(最大ベクトル性能)を持つ最新鋭のベクトル並列型スーパーコンピュータです。本サーバは 5 つのノードから構成され、それらは高性能なネットワーク装置(IXS、ノード間双方向バンド幅 32GB/s)で相互接続されています。本サーバは 5 ノード全体を利用者に一括提供することを考えており、これにより最大ベクトル性能 640GFLOPS、メモリ容量 640GB の大規模、かつ高性能な処理が可能になります。現在運用中の SX-7 では、1 ノードを占有して最大 256GB メモリと 282GFLOPS の演算性能を利用することが 1 つのプログラムの実行の上限となっておりますが、ベクトル並列型演算サーバの 5 ノード全体で提供される 640GFLOPS の性能、および 640GB のメモリ容量を必要とするような大規模計算の実行をご希望の利用者の方は、センターのシステム管理係にぜひご相談ください。ノード間並列処理には MPI を使った並列プログラミングが必要となりますが、5 ノードを使った並列処理において最大性能を引き出すための高速化支援を用意しております。

オペレーティングシステムは、SX-7 と同じ SUPER-UX であり、表 3 に示すように SX-7 の利用環境と同じものを提供しております。

表 3 ベクトル並列型演算サーバで利用可能なコンパイラおよびライブラリ

FORTRAN90/SX	ISO/IEC 1539-1:1997,自動並列化, OpenMP
C++/SX	ISO/IEC 14882:1998,自動並列化, OpenMP
MPI/SX	メッセージパッシングライブラリ
ASL	数値計算ライブラリ

(3) ファイルサーバ

ファイルサーバは、スカラ並列型演算サーバとベクトル並列型演算サーバにファイバチャネルで接続され、GFS(グローバルファイルシステム)による高速共有ファイルシステムを提供しています。これにより、各サーバ間の高速で一貫したデータ共有を可能とし、スカラ並列型演算サーバとベクトル並列型演算サーバを連携させた計算において、より一層の利用効率の向上を実現しています。なお、スーパーコンピュータSX-7との間は従来通りNFSによるファイル共有です。

ファイルサーバは、10 テラバイト(RAID5 時)の記憶容量を持つ磁気ディスク装置(iStorage S2400、日本電気(株)製)と、15 テラバイトの記憶容量を持つテープバックアップ装置(iStorageT40A、日本電気(株)製)から構成され、安心・安全・高性能なファイルシステムを利用者に提供いたします。

(4) その他

バッチ管理システムとして NQSII(Network Queuing System II)を導入し、スカラ並列型演算サーバ(TX7/i9610)、ベクトル並列型演算サーバ(SX-7C)、現有スーパーコンピュータ SX-7 のバッチジョブの管理を一元化し、利便性の高いジョブ投入環境を提供します。さらに、SX-7 も含めた全システムの一元的な利用者管理と課金統計管理を行います。なお、新並列コンピュータシステムの利用負担金については現在検討中です。

3. 移行日程

現システム(TX7/AzusA)は、2006 年 3 月 16 日(木)で運用を終了する予定です。新並列コンピュータシステムは、2006 年 4 月 3 日(月)からの運用を予定しておりますが、それに先立って、3 月 22 日(水)からの試験運用を計画しております。ぜひ、最新の Itanium2 プロセッサと大規模ベクトル並列処理の能力を試してみてください。なお、並列コンピュータのサービスは、3 月 17 日から 3 月 21 日の 5 日間、全面停止の予定です。利用者の皆様のご理解とご協力をお願いいたします。並列コンピュータの正式の停止日も含めて、新システムの詳細については、本広報誌 SENAC、および「大規模科学計算システムニュース」で順次お知らせする予定です。不明な点、ご質問等

がございましたら、お気軽に本センターの共同利用係へお問い合わせください。



スカラ並列型演算サーバ(TX7/i9610 1 ノード)・ファイルサーバ



ベクトル並列型演算サーバ(SX-7C 5 ノード)